

¹Application of Point Density Estimation in Assessing Changes in Polarization in Tanzania between 1992 and 2001

John K. Mduma²
Department of Economics
University of Dar es Salaam

April 25, 2008

Abstract

Data from the recent two Household Budget Surveys of 1991/1992 and 2000/2001 in Tanzania indicate that there is no change in the inequality between the two surveys. Yet there is increasing discontent throughout the country that the economic reforms from socialist to market economy in the 1990, although registered impressive macroeconomic indicators, worsened the income distribution. We argue that the Gini index failed to capture other inconspicuous trends in the income distribution, particularly the problem of polarization across space (regions). Using polarization measure based on a point density estimation of alienation and identification, this study analyzes changes in the distribution of household income in Tanzania in the 1990s. The study finds that, unlike the Gini index, polarization index increased significantly between 1992 and 2001. The study also finds evidences of increased spatial variability across regions and lack of spatial convergence in household incomes. Policy implications are also discussed.

1. Motivation and Introduction

Data from the recent two Household Budget Surveys (HBS) of 1991/1992 and 2000/2001 in Tanzania indicate that there is no change in the inequality between the two surveys ([Demombynes and Hoogeveen 2004](#)). Yet there are increasing discontents throughout the country that the economic reforms in the 1990 from socialist to market economy, although registered impressive macroeconomic aggregates, worsened the income distribution. It is not uncommon to hear anecdotal comments that the reforms have only benefited few at the expense of the majority. Although the Gini indexes between 1991 and 2001 do not indicate significant changes in the overall income inequality, it does not rule out the possibility that other inconspicuous trends in the income distribution, particularly the problem of polarization across space (regions) and over time has actually increased.

Policy issues which this study addresses are: Does the Gini index tell the entire story about the changes in income distribution in the 1990s? If yes, why is there loud outcry that the reforms of the 1990s have resulted in ill-distribution in the income? Is this outcry founded on some aspects of the distribution that the Gini index fails to capture? What is the spatial variability of this index?

Tapping from the recent theoretical and estimation developments, we analyze the spatial-temporal variations in polarization in Tanzania using the point density estimation of Identification-Alienation Framework proposed in [Duclos et al. \(2004\)](#). The main premise of this framework (which differs from the standard inequality measurement) asserts that

¹ A paper prepared for the Eleventh Annual Conference on Econometric Modeling in Africa

² e-contact: nashmduma@yahoo.com

alienation must also be complemented by a sense of identification such that a combination of the two generates an index which is sensitive (in the same direction) to both elements of inequality and equality. This last attribute of polarization index is what Gini index fails to capture.

One of the main econometric problems in the estimation of the polarization indexes is how to estimate the size of the groups to which individuals belong ([Wolfson 1994](#); [Esteban and Ray 1994](#)). It is noted that the Esteban-Ray polarization index is based on a discrete and finite set of income groupings located in a continuous income distribution. As a result, two problems arise. The first one is the artificial discontinuity introduced by grouping of a continuous distribution. The second problem is a working assumption that the population has already been bunched in the relevant groups. This last one is more serious for the present study because there is no reason to assume that the 'relevant groups' have remained stable throughout the 1990s-when Tanzania witnessed massive social and economic transformation (URT 2002). To circumvent both problems associated with the use of Esteban-Ray polarization index, the paper is anchored on the framework of [Duclos *et al.* \(2004\)](#). In this case, group based-estimation is replaced with point density estimates, whereby the density is estimated non-parametrically.

Data for this study come from the Tanzanian HBS of 1991/1992 and 2000/2001 and Tanzania-GIS data from the International Food Policy Research Institute (IFPRI). A brief description of the GIS data can be found in [Mduma and Wobst \(2005\)](#). The empirical analysis establishes that there has been a statistically significant increase in polarization in Tanzania between 1992 and 2001. The increase has also been shown to have substantial spatial variability but only few regions in the country registered a decline in polarization.

The rest of the paper is organized as follows. Section 2 provides a brief overview of Tanzania income distribution. Section 3 presents the theoretical and empirical framework used in this study. The results of the analysis and discussions are presented in Section 4. Section 5 concludes.

2. Overview of Income Distribution in Tanzania

Since 1967 to the second half of 1980s Tanzanian economy was managed under socialist taboo, where excessive income inequality was strongly discouraged³. Thus, income inequality in Tanzania was less pronounced when compared to most of Tanzania's neighbors. Under socialism, most large companies were managed by the government. As such the government became the (sole) largest employer both in the urban and in the rural areas (outside small scale agriculture) whereby it standardized most salaries scales.

In the area of enterprising, there were no incentives for the expansion of private entrepreneurship because it was considered exploitative. Furthermore, rural cooperatives, crop boards, and the regional trading companies were the only legal institutions allowed to

³ Prior to 1967, the income inequality was relatively high but was radically reduced after the introduction of socialist policies (which entailed mass nationalization of command heights of the economy) in 1967.

carry out large scale commercial activities in rural areas. As a result of these measures wide inequality in the income distribution was less common in various sectors of the population.

As the 1980s closed up, Tanzania's socioeconomic orientation started to drift away from socialist principles and gradually embarrassing the market economy policies. In the early phases of the reforms, important policies included licensing private crop buyers, abolishing government control over the cooperatives, and the privatization of the state parastatals.

These reforms were further intensified in the 1990s where Tanzania enacted several laws that unleashed the market forces. As such, the nationally (and regionally) representative HBS in 1991/92 was conducted at the time when Tanzania was undertaking wide spread reforms in several fronts of the economy. The timing of the survey is such that it forms the baseline data on which impacts of subsequent economic reforms can be assessed.

The following HBS was conducted in 2000/2001 and by comparing the two we can assess what happened in the income distribution in the 1990s. The results of this comparison are summarized in Table 1, where it is apparent that there have been relatively small changes in the income distribution (using these standard indicators). But as we shall show soon, these indicators do not tell the whole story.

Table 1: Consumption Data from Household Budget Survey Data

	Mainland Tanzania	Dar es Salaam	Other Urban areas	Rural areas
<u>Consumption per Adult Equivalent</u>				
Mean, 1991/92	10223 (290)	11161 (410)	12445 (842)	9820 (326)
Mean, 2000/01	10880 (250)	15944 (779)	13536 (487)	10060 (273)
Ratio (00/01) to (91/92)	1.064	1.429	1.088	1.024
<u>Gini Coefficient</u>				
1991/92	0.33 (0.01)	0.30 (0.01)	0.34 (0.02)	0.33 (0.02)
2000/01	0.34 (0.01)	0.34 (0.01)	0.35 (0.01)	0.32 (0.01)

Note: standard error in the brackets

Source: Demombynes and Hoogeveen (2004)

Table 1 also shows further that the rise in household consumption was accompanied by a small increase in inequality. The Gini coefficient, a measure of how unequal expenditure is distributed, rose marginally from 0.33 to 0.34. Using quintiles of the income distribution, it is found that the richest 20 percent of the population now account for 44 per cent of household spending, compared with 43 per cent in 1991/92, which again is just a marginal increase (NBS 2002).

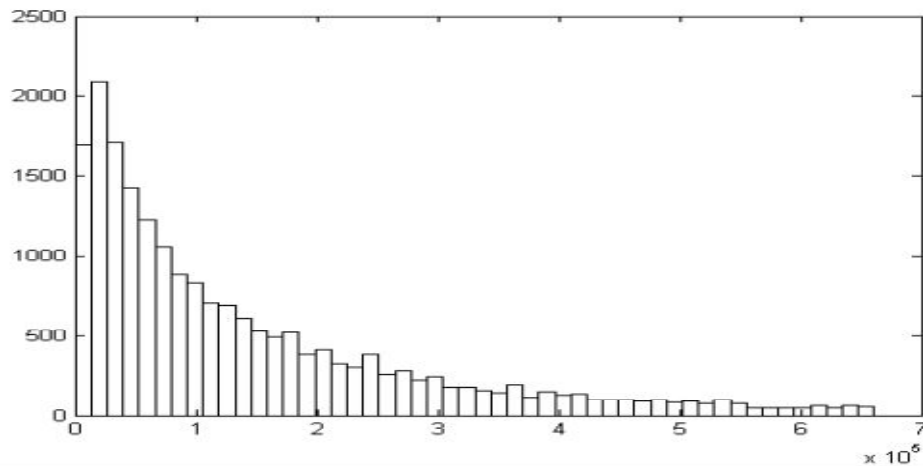


Figure 1a: Distribution of per capita expenditure in 2001 in Tanzania

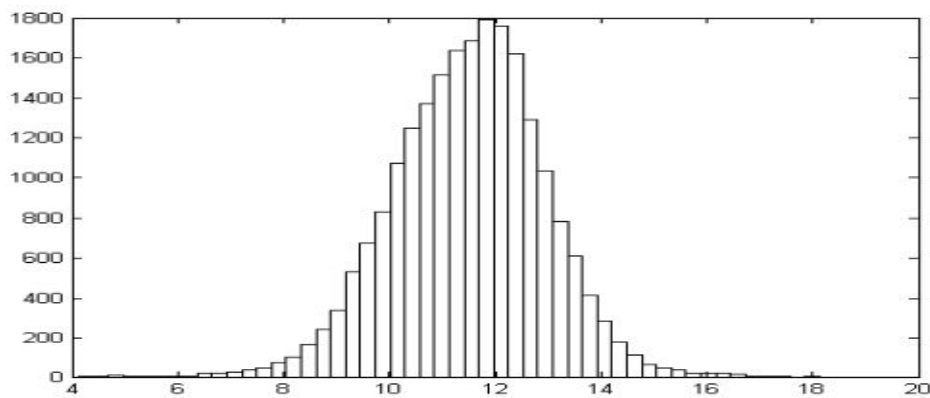


Figure 1b: The distribution of log of per capita expenditure in 2001 in Tanzania

As expected (since income tends to follow a lognormal distribution), the median – which shows the value below which half of the population falls – was substantially lower than the mean, because a small number of high values have more effect on the mean (Figure 1a & 1b). Furthermore, earlier studies which use 2000/01 HBS data have also revealed that the distribution of per capita expenditure data also shows substantial spatial variability (see Mduma and Wobst 2005, NBS 2002). For example, the mean per capita expenditure in Dar es Salaam is 2.6 times higher than in rural areas. NBS (2002) argue that the differences in average consumption expenditure per household were slightly smaller in the rural areas because rural households tend to be larger.

By comparing the per capita expenditure across regions, it is found that the average per capita expenditure is highest in Dar es Salaam, Mbeya and Mtwara and lowest in Rukwa, Singida and Kigoma. It is straight forward to explain why there is relatively high per capita income in Dar es Salaam and Mbeya regions (for Dar es Salaam is the commercial capital of Tanzania and Mbeya is fairly favored agro-ecologically). However, there is no good explanation for Mtwara region which is known to perform poorly in other welfare indicator (NBS 2002). In the next section we offer theoretical and empirical strategies of measuring polarization index used to explore other aspects of income distribution beyond the Gini index.

3. Theoretical and Empirical Strategies

Measurement of income polarization, in theory and practice, can be considered as an extension of that of income inequality. The difference is that polarization is related to the alienation that individuals and groups feel from one another whereby such alienation is fuelled by notions of within-group identity (Duclos *et al.* 2004). In its broad sense, the concept of polarization has been used in describing the diminution of the middle class in wage, income and wealth distributions, in studying growth and convergence issues, in the studies of inter-generational income relationships, and in discriminating between competing matching models of marriage partners. Wolfson (1994) and Wolfson (1997) are some of the studies which provide critical review of the various domains where the concept of polarization has been applied.

Prior to Duclos *et al.* (2004), the literatures broadly interpret polarization as the disappearance of mass at the center of income distribution or the increasing distance between multiple points of modality of the distribution. In this early literature, the index of polarization followed the estimators proposed by [Esteban and Ray \(1994\)](#). This index is based on several axioms whereby polarization is jointly determined by an agents' sense of increasing within group identity and between group distances. Formally, the Esteban-Ray polarization index is given by:

$$P_{\alpha} = K \sum_{j=1}^n \sum_{i=1}^n (|x_j - x_i| \pi_i^{1+\alpha} \pi_j) \quad (1)$$

where K is a normalizing constant, π_i is the sample weight of the i^{th} observation and where $\alpha \geq 0$ is a polarization sensitivity factor chosen by the analyst.

Empirical operationalization of equation (1) perceives the distribution of income as a mixture of sub-distributions. As such, this thinking is the pillar of the initial formulation of Esteban and Ray (1994) polarization index; thereby the population is discretely divided into finite set of income groupings. Thus the estimator in (1) presumes that such groups can be optimally established, which is difficult to justify in this exploratory study where we explore what has happened in the income distribution between 1991 and 2001. As discussed in Section 2, reforms implemented in the 1990s are likely to have had substantial changes in the configuration of the income groups required to implement the estimator in (1).

To explore changes in the configuration of income groups that occurred in the 1990s, we first estimate Adaptive Finite Mixture (AFM) models to determine what would be the mixtures in the distributions of household income in years 1991 and 2001. We initially start with five mixtures and incrementally increased the search to 25 mixtures. For the 2000/01 HBS expenditure data, it was established that there are eight optimal mixtures. For 1991/91 HBS expenditure data, the optimal number of groups happened to be 23. The results of this analysis are summarized in Annex Figure A1.

From the results of Adaptive Finite Mixture, it is evident that polarization index based on the discretization of the distribution of the expenditure can be quite misleading due to heterogeneity in the configuration of income groupings between 1992 and 2001. The difficulties caused by the heterogeneity in the mixtures between the two dataset can be overcome by the use of polarization measure which does not depend on the discretization of the distributions. Such measure has recently been suggested by Duclos *et al.* (2004) where identification is based on the point density estimation.

The point density approach is linked to the discretization based approach in that individuals possess a “window of identification” with weights decreasing with the distance. At the same time, individuals would feel alienated only from those outside the window. The major departure of the former from the later is that the point density estimation is a limit case as the window of identification is composed of individuals having exactly the same income. This estimator is given as

$$P_{\alpha}(F) = \int_y f(y)^{\alpha} \int_x |y - x| dF(y) dF(x) \quad (2)$$

where, as in the case of polarization based in discrete estimator, α is the polarization sensitivity factor, $F(y)$ is a CDF and $f(y)$ is a PDF. Duclos *et al.* (2004) based on several axioms, shows that for a meaning index of polarization in this framework, $\alpha \in [0.25, 1]$.

It can be noted from equation (2) that polarization is an interplay of two forces. The first one is the identification with one’s own group and the second one is the alienation vis-à-vis others. It should also be noted that, in general, these two forces are inseparable because they are jointly explained by the same density. However, for analytical purpose, the final amount of polarization depends on the separate contributions of alienation, identification, and their covariance. From this discussion, an α -identification at income y , denoted by τ_{α} , is measured by $f(y)^{\alpha}$ so that the average α -identification $\bar{\tau}_{\alpha}$ is defined by:

$$\bar{\tau}_{\alpha} = \int f(y)^{\alpha} dF(y) = \int f(y)^{1+\alpha} dy \quad (3)$$

On the other component, the alienation between two individuals with incomes y and x is given by $|y - x|$. To get the overall alienation felt by an individual with income y , one has to integrate over the support of x , formally expressed as:

$$a(y) = \int |y - x| dF(x) \quad (4)$$

From the individual level of alienation, the average alienation, over the population is obtained by integrating (4); expressed as:

$$\bar{a} = \int a(y) dF(y) = \int |y - x| dF(x)dF(y) \quad (5)$$

If we define ρ as the normalized covariance between identification and alienation, Duclos *et al.* (2004) shows that the polarization index can be expressed as:

$$P_\alpha(f) = \bar{a}\bar{\tau}_\alpha (1 + \rho) \quad (6)$$

For this study, $F(x)$ is estimated non-parametrically, using Gaussian kernel. As usual the distribution of the expenditure data is skewed (arguably, it follows lognormal distribution, see Figure 1). Thus, in the kernel density estimation, we use the bandwidth proposed by [Hardle \(1990\)](#) in order to obtain robust estimator of the density. Furthermore, it has been shown that the estimator in (6) has a limiting Gaussian distribution. Thus, to decide whether there are significant changes in the polarization, we use the t-test of the equality of the two means under the assumption of independence between the two samples (surveys) and unequal variances ([Anderson 2005](#)).

So far the discussion in this section has dwelt on the strategy of estimating the polarization index. The strategy so far discussed can only tell us the extent of polarization in the two survey years and whether there has been significant change in polarization between the two surveys. To investigate the spatial pattern of polarization, we resort to Spatial Econometrics. Important questions addressed using this technique are whether or not there is spatial correlation in the polarization index in 1992 and 2001, and whether or not changes in Gini and polarization index in the 1990s show spatial co-variation. With this, spatial autocorrelation in the estimated index is explored using the First Order spatial Autoregressive Model (FAR), which is specified as:

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \boldsymbol{\varepsilon}; \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}) \quad (7)$$

where \mathbf{y} is a vector of index of our interest (in this case, regional polarization index), $\boldsymbol{\varepsilon}$ is a vector of spatial error Gaussian process, and \mathbf{W} is n-by-n matrix of spatial weights.⁴ To explore the spatial correlation in two variables namely changes in the polarization and Gini index between 1992 and 2001, we use Spatial Error Model (SAR)⁵, which is specified as:

⁴ The \mathbf{W} matrix is produced using Delaunay Triangulation. Given a set of coplanar data points (points that lie on the same plane), Delaunay triangulation is a set of lines connecting each point to its natural neighbors. It is a set of triangles such that no data points are contained in any triangle's circumcircle ([Haining 1993](#)).

⁵ For this, simple Moran I test or FAR model could not be used because it uses only one variable, thus we resort to using Spatial Autoregressive Model.

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \beta\mathbf{x} + \boldsymbol{\varepsilon}; \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}) \quad (8)$$

where \mathbf{y} is a vector of changes regional polarization index and \mathbf{x} is the second variable of interest, in this a vector of changes in the regional Gini index; and other notations are as defined earlier. We used changes in the polarization index as the dependent variable and changes in the Gini index as the explanatory variable to capture the sense of correlation between the two.

4. Results and Discussion⁶

Since this paper is an exploratory study, the discussion of the results started much earlier, when we presented the general overview of income distribution in Tanzania and the results of the Adaptive Mixture Models (used to justify the use of point density estimation of the polarization index). In this section, we continue with the discussion of the results of the estimated polarization indexes, first at the national level, the rural urban levels, and then at the regional levels. We first discuss results pertaining to the national and rural-urban levels. It is worth noting at this stage that the results also show substantial sensitivity to the parameter of alienation α . However, since we observed a consistent pattern after using several values of α in $[0.25, 1]$, we only discuss the results of polarization when $\alpha = 0.5$.

4.1. Polarization at National Level and Rural and Urban Areas

The analysis indicates that there is polarization in the whole sample which was statistically different from zero at one percent in both years 1992 and 2001. In particular, our results reveal that the overall level of polarization in 1991 was about 0.28747284 (with standard error of 0.01742980), which by 2001 it has increased to 0.37859119 (with standard error of 0.01533438). This increase is statistically significant at 1% level (empirical z-value was 21.9445). This finding was robust irrespective of the value of α chosen⁷. Comparatively, the results indicate that there is substantial polarization in Tanzania in 2001 when gauged against polarization reported in OECD countries (Duclos *et al.* 2004). Unfortunately, there are no similar studies from developing countries to compare.

From these results, we agree that even though the Gini index did not show significant change in the income inequality between 1991 and 2001, the complaints and the outcry about ill-distribution of incomes in the 1990s are supported by the empirical evidence (in terms of polarization). Thus, the results of this study are a wake-up call that policy to address issues of income distribution needs to look beyond the standard measures such as the Gini index.

Table 2 provides further information on the changes in income distribution in Tanzania. It is observed that, overall polarization increased by around one-third between 1992 and 2001 and that most of the increase is urban driven. As such, the increase in polarization in the

⁶ Polarization and Gini indexes were estimated using DAD software (Duclos et al. 2004) and the spatial models were estimated using MATLAB Econometrics Toolbox (LeSage 1999).

⁷ Since $\alpha \in [0.25, 1]$, we verified this assertion by taking few values of α , including its boundary values.

rural areas was not statistically significant. This implies that the economic opportunities or constraint introduced by the reforms in the rural sector has not been in favor/against distinct groups in the rural areas. The contrary can be said for the urban areas.

Table 2: National, Urban and Rural Outlook of Polarization in Tanzania

2001	Polarization Index	Alienation: $E(a(y))$	Identificación $E(f(y)^\alpha)$	Normalized Covariance: ρ
National	0.378591	0.692031	0.798401	-0.31478
Urban	0.369034	0.680033	0.793044	-0.31571
Rural	0.330234	0.609965	0.731645	-0.26003
	% change from 1991			
National	31.69633	-14.5041	42.23347	-14.2979
Urban	29.83127	3.120483	31.15936	9.950704
Rural	1.640286	-26.8947	16.79855	-31.2771

Source: Author estimation from Tanzania's HBS data for 2001 and 1991

Table 2 also reports the average alienation, average identification, and the normalized covariance between the alienation and identification. When looking at the changes in covariance of alienation and identification, it is apparent that there has been a decrease in their co-movement except in the urban area. For the urban areas, it can be interpreted that alienation and identification reinforced each other in explaining the increase in polarization.

Furthermore, it is apparent in Table 2 that there is more alienation in the urban areas than in the rural areas. Comparison between rural and national average indicates that there is less polarization in the rural areas than the national average. Further decomposition of the polarization index indicates that there were larger changes in the identification than in the alienation, which corroborate well with the earlier findings that there was not substantial changes in the Gini Index (NBS 2002). As such, this study tends to suggest that much of the problems lie on the identification side than in the actual alienation.

4.2. Further Spatial-Temporal Variability

The discussion of the rural-urban analysis in the preceding subsection already indicates that there are likely to be substantial spatial variations in the polarization between 1991 and 2001. This subsection discusses the results of the analysis of polarization index and Gini index computed spatially at regional level for year 1991 and 2001.

Table 3: Summary of the Regional Indicators of Polarization in Tanzania

Indicator	Mean	Std. Dev.	Min	Max
	2001			
Polarization index	0.387006	0.082463	0.303557	0.667262
Alienation: $E(a(y))$	0.647028	0.086274	0.548356	0.911002
Identification: $E(f(y)\alpha)$	0.818191	0.10551	0.730373	1.197008
Normalized Covariance: ρ	-0.26991	0.043918	-0.3881	-0.17545
	1991			
Polarization index	0.350028	0.052171	0.271277	0.503046
Alienation: $E(a(y))$	0.681649	0.120787	0.513393	0.882341
Identification: $E(f(y)\alpha)$	0.72293	0.084124	0.541701	0.858124
Normalized Covariance: ρ	-0.27552	0.090281	-0.41695	-0.12622

Source: Author's estimations from Tanzania's HBS data for 1991 and 2001

Table 3 presents the summary statistics of polarization and its components from the regional estimates where it is apparent that there are substantial variations in polarization index and its components. Further analysis of the variability indicated that the coefficient of variation was 21.3% for the polarization index in 2001 as compared to 14.9% in 1991. The coefficient of variation of identification increased marginally from 11.5% to 12.9% in the same period. However, the variability of alienation was higher in 1991 than in 2001. In general, the increase in mean polarization and its variability indicates divergence across regions in the 1990s. Since the study uses per capita expenditure, which serves as an important indicator of welfare, our results can also be interpreted to imply that there is also an increase in regional welfare disparities.

Another aspect analyzed spatially is the co-variation in the change in the regional polarization and Gini indexes between 1991 and 2001. We find that there is significant correlation between the changes in polarization index and the changes in the Gini coefficient. Spearman rank correlation coefficient was 0.598 (p-value = .005 for two tail test). However, Annex Figure A1 shows that the two indexes changed in the same direction only in 12 out of 20 regions of Tanzania. We also note that there are four regions in which both the polarization index and the Gini index declined in the 1990s. Thus, this exploratory analysis shows the importance of carrying the analysis at lower levels to reveal the information that would otherwise be concealed by a national level analysis.

So far the discussion of our results has focused on what would be termed 'quasi-spatial' because, the discussion has only been captured in sense of regional variations. Limiting our discussion there would not be able to reveal whether or not geographically neighboring regions tends to be characterized by similar pattern in the changes in the polarization. However, spatial econometric analysis interesting results, which could otherwise go unnoticed. In particular, the results of the First Order Spatial Autoregressive (FAR) Model in Table 4 show that there is significant spatial correlation in polarization in both 1992 and

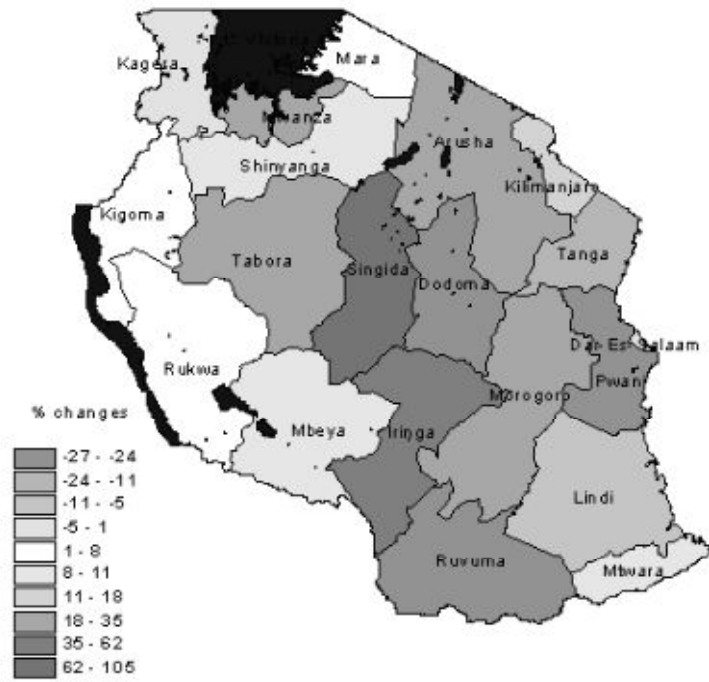
2001. However, the changes in the polarization index do not show statistically significant pattern. Likewise, the changes in Gini indexes between the two years do not show systematic spatial pattern. These results once again indicated the lack of spatial convergences in income distribution. In other words, what happened to the income distribution of one region did not necessary have effects on the income distribution of the neighboring regions. The lack of such pattern will imply the lack of spatial multiplier (Anselin 2002).

Table 4: Estimate of First Order Spatial Autoregressive of Polarization

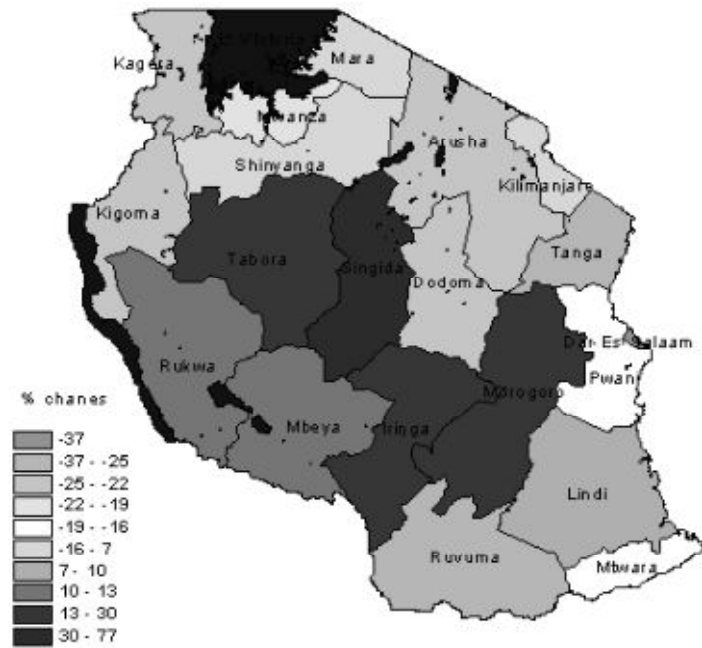
Variable	Coefficient	Asymptot t-stat	t-probability
Pola 1991	0.955978	24.727568	0.000000
Pola 2001	0.904970	11.124400	0.000000
Polar change	0.336955	0.843409	0.399000
Gini change	-0.174988	-0.327713	0.743129

Source: Author's own computation from Tanzania's HBS data for 1991 and 2001

The spatial analysis was extended further beyond the FAR estimation to Spatial Autoregressive (SAR) estimation in order to assess the spatial pattern (correlation) between changes in the Gini index and the index of polarization. The coefficient of association was found to be 0.836113 (t-value = 4.708901), indicating significant association between changes in the two index. Figure 1 provides a spatial visualization of this correlation. Annex Figure A3 further shows this spatial co-variation whereby conditioning on the positive change in the polarization index, we observed relatively large value of changes in the Gini index. This is evidenced by a leftward shift and a sharp peak in a conditional density of changes in Gini index (conditional on positive changes in the polarization index).



Polarization



Inequality

Figure 1: Spatial Distribution of the Changes in the Polarization and Gini Indexes in Tanzania Between 1991 and 2001

Note that Figure 1 is composed of two maps of Tanzania, indicating changes in the regional index of polarization and Gini index, respectively. It is apparent that there is substantial spatial match in the two changes. For both indexes, the maps show that larger positive changes are concentrated in the middle of the country. For the case of changes in Inequality, Ruvuma region in the south happens to be a spatial outlier, registering huge decrease as compare to most of its neighbors. In terms of polarization, Kagera and Dodoma regions happens to be spatial outlier. Note that regions which have been associated with emigration in the 1990s happen to experience relatively high level of polarization but not necessary high inequality (Mduma and Wobst 2005).

5. Summary and Conclusions

This study aimed at investigating changes in the income distribution in Tanzania between 1992 and 2001 and across space. The study was motivated from the view that there has been insignificant increase in the income inequality as measured by the Gini index while at the same time there has been a loud outcry that the economic reforms of the 1990s have resulted in ill-distribution of incomes.

Combining a number of analytical tools, ranging from Adaptive Finite Mixture, Kernel Density Estimation, and Spatial Econometrics, the study has demonstrated that there is a significant increase in the polarization in Tanzania between 1991 and 2001, which was not necessary matched with the increase in inequality. The spatial analysis indicated that there is clearly defined spatial pattern in the polarization index in both 1991 and 2001 such that regions with high polarization show spatial clustering. However, the rate at which this index has changed between the two reference years does not exhibit any systematic spatial patterns. Nevertheless, the rate at which polarization and inequality has changes was found to have strong spatial correlation.

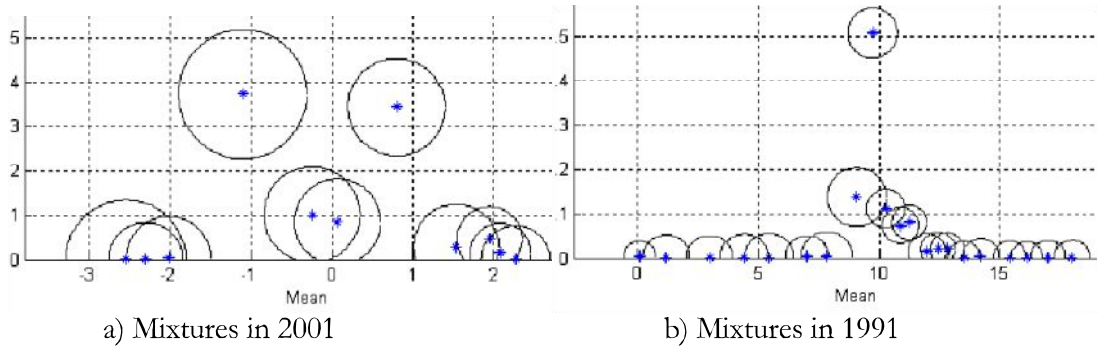
The study not only shows the increase in polarization across regions of Tanzania but also the increase in the variability of polarizations. In line with the literature on spatial convergence, this phenomenon was interpreted to imply that in 1990s Tanzania recorded increase in regional disparities, at least in terms of income distribution. The increase in polarization and lack of spatial convergence are issues that should be addressed to avoid social disorders.

References

- [Anderson G.J. \(2005\). "Indices and Tests for Alienation Based Upon Gini Type and Distributional Overlap Measures." University of Toronto Economics Department.](#)
- [Anselin, L. \(1988\). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.](#)
- Anselin, L. (2002). "Spatial Externalities". Forthcoming in *International Regional Science Review*. Online manuscript at: www.agec221.agecon.uiuc.edu/users/anselin/papers/intro_externalities.pdf; accessed April 29, 2004.

- Barry, R and R. Kelley Pace (1999). "A Monte Carlo Estimator of the Log Determinant of Large Sparse Matrices". *Linear Algebra and its Applications*, Vol. 289, No. 1-3, pp. 41-54.
- Demombynes, G and J.G. Hoogeveen (2004). "Growth, Inequality and Simulated Poverty Paths for Tanzania, 1992-2002". World Bank Policy Research Working Paper 3432, October 2004.
- Duclos, J.Y., J. Esteban and D. Ray (2004). "Polarization: Concepts, Measurement, Estimation". *Econometrica*, Vol. 72, No. 6, pp. 1737-1772
- Duclos, J.Y, A. Araar and C. Fortin (2004). DAD: Distributive Analysis User's Manual. Université Laval <http://www.mimap.ecn.ulaval.ca/>
- Esteban J., and D. Ray (1994). "On the Measurement of Polarization". *Econometrica*, 62, 819-852.
- Haining, R. (1993). *Spatial data analysis in the social and environmental sciences*. New York: Cambridge University Press
- Hardle, W. (1990). *Applied Nonparametric Regression*. Cambridge: Cambridge University Press.
- LeSage, J. (1999). Spatial Econometrics, Department of Economics, University of Toledo. <http://www.spatial-econometrics.com/>; accessed April 1, 2001.
- Mduma, J. and P. Wobst (2005) Village Level Labor Market Development in Tanzania: Evidence from Spatial Econometrics. ZEF-Discussion Papers on Development Policy 96, Center for Development Research, Bonn.
- NBS (2002). *Household Budget Survey 2000/01*. Tanzania National Bureau of Statistics, Dar es Salaam.
- Pace, R. K. and Ronald P. Barry (1998). "Simulating Mixed Regressive Spatially autoregressive Estimators". *Computational Statistics*, Vol. 13, pp. 397-418.
- URT. *Economic Survey* (various). Dar es Salaam: National Planning Commission.
- Wolfson M.C. (1997). "Divergent Inequalities: Theory and Empirical Results". *Review of Income and Wealth*, Vol. 43 pp. 353-358.
- Wolfson, M.C. (1994). "When Inequalities Diverge". *American Economic Review Papers and Proceedings* Vol. 84, pp. 353-358.

Annexes



Note: Vertical axis measures the mixing coefficient and horizontal axis measures group means in the normalized scale.

Figure A1: Comparison of the Income Groups Generated by Adaptive Finite Mixture for HBS Data of 1991 and 2001.

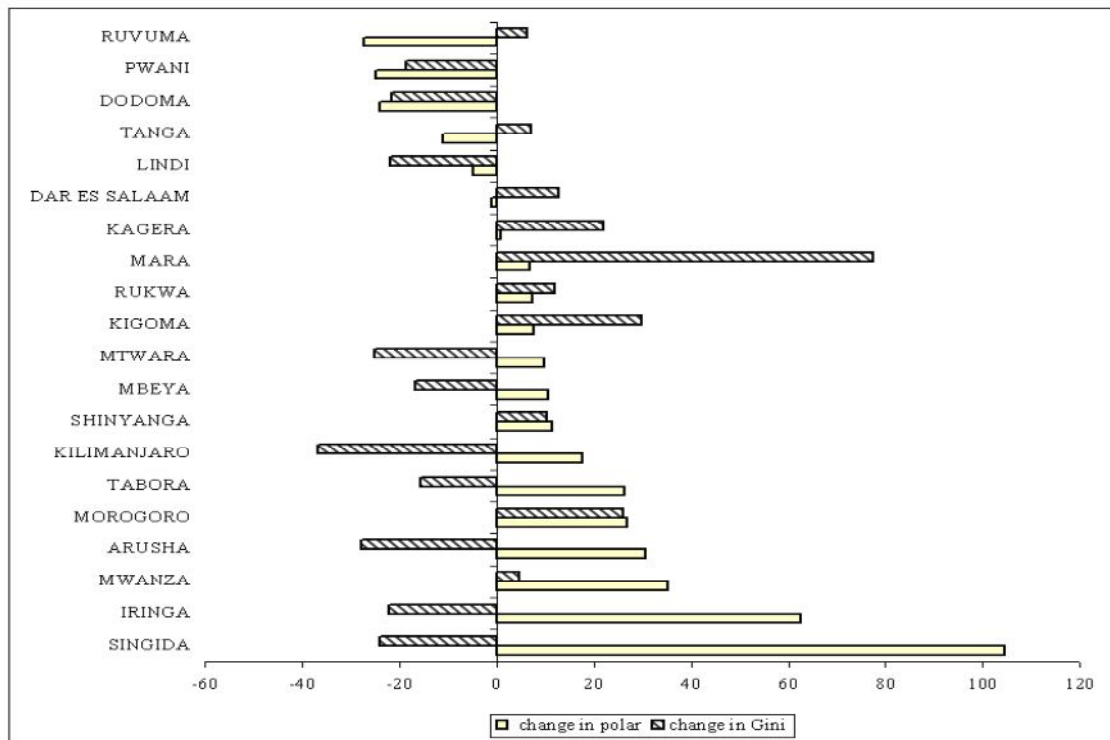


Figure A2: Change in the Regional Polarization and Gini Indexes between 1991 and 2001

Table A1: Estimated Regional Polarization indexes for 2001

Regions	Estimate (DER)	Alienation: $E(a(y))$	Identification: $E(f(y)^2)$	Normalized Covariance: ρ
DODOMA	0.30985254	0.55089206	0.76104372	-0.26093976
ARUSHA	0.35421359	0.59973950	0.80568603	-0.26694431
KILIMANJARO	0.35978022	0.60353429	0.79191770	-0.24724254
TANGA	0.33425221	0.56849045	0.78622580	-0.25216804
MOROGORO	0.45125555	0.74332735	0.89440217	-0.32125023
PWANI	0.37734745	0.65043954	0.79337250	-0.26876266
DAR ES SALAAM	0.31828721	0.55749629	0.73037259	-0.21831388
LINDI	0.36865475	0.62766281	0.81340255	-0.27791450
MTWARA	0.35177046	0.66326186	0.74171165	-0.28494584
RUVUMA	0.30355684	0.54835546	0.76239746	-0.27390088
IRINGA	0.51934969	0.76396212	0.95502456	-0.28817729
MBEYA	0.36261139	0.57696950	0.76220454	-0.17545074
SINGIDA	0.66726218	0.91100173	1.19700753	-0.38809858
TABORA	0.39847410	0.65320551	0.80958705	-0.24649508
RUKWA	0.35181473	0.61517032	0.74565710	-0.23303056
KIGOMA	0.36970331	0.64796714	0.81075109	-0.29626119
SHINYANGA	0.41016773	0.69177913	0.87746056	-0.32428260
KAGERA	0.34087623	0.61454255	0.73196515	-0.24220169
MWANZA	0.39121111	0.66898309	0.81481004	-0.28230460
MARA	0.39968309	0.68376940	0.77881027	-0.24945968

Table A2: Estimated Regional Polarization indexes for 2001

Regions	Estimate (DER)	Alienation: $E(a(y))$	Identification: $E(f(y)^2)$	Normalized Covariance: ρ
DODOMA	0.40806275	0.72495665	0.80365055	-0.29959681
ARUSHA	0.27127714	0.77036093	0.54170131	-0.34993258
KILIMANJARO	0.30606714	0.57679294	0.70508154	-0.24741107
TANGA	0.37564545	0.78779157	0.75766453	-0.37065335
MOROGORO	0.35577168	0.59063240	0.72863431	-0.17330673
PWANI	0.50304613	0.77167711	0.85812433	-0.24033516
DAR ES SALAAM	0.32219202	0.88234071	0.62628784	-0.41695260
LINDI	0.38742131	0.56927254	0.77885900	-0.12621559
MTWARA	0.32081080	0.79790861	0.63426913	-0.36609836
RUVUMA	0.41846174	0.73255570	0.85000363	-0.32796148
IRINGA	0.31962876	0.58891626	0.68706786	-0.21006230
MBEYA	0.32831405	0.51556477	0.76609283	-0.16876315
SINGIDA	0.32627689	0.51339293	0.74964341	-0.15222380
TABORA	0.31558232	0.53531394	0.72426542	-0.18603376
RUKWA	0.32779767	0.54614915	0.77708799	-0.22763232
KIGOMA	0.34351463	0.83089100	0.66663212	-0.37982362
SHINYANGA	0.36846547	0.64707624	0.76564538	-0.25627315
KAGERA	0.33850388	0.78423594	0.68268600	-0.36774034
MWANZA	0.28953314	0.82340178	0.57447589	-0.38790983
MARA	0.37418444	0.64375762	0.78072069	-0.25549475

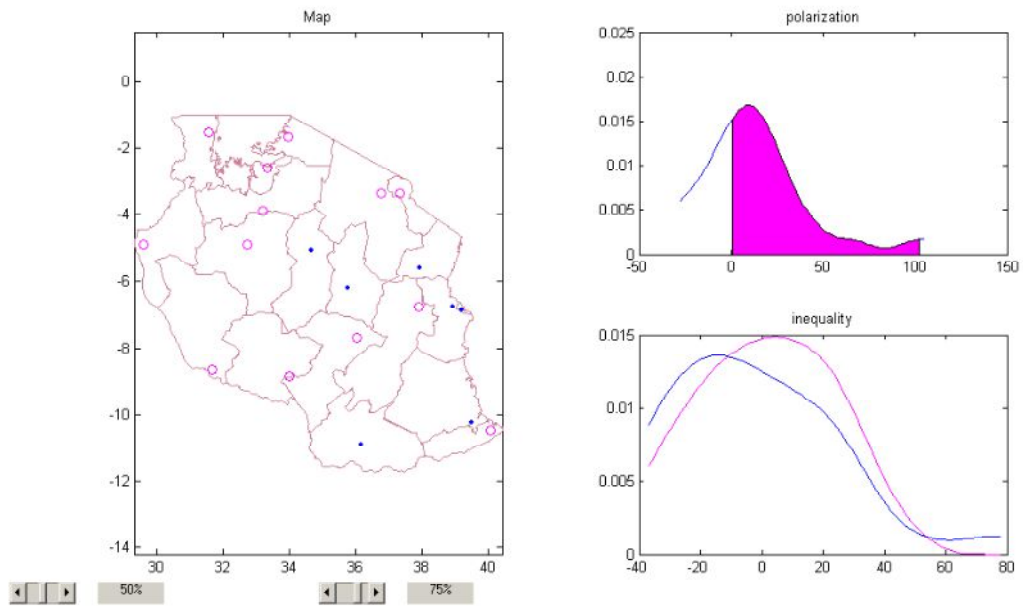


Figure A3: Further evidence of spatial association of between change in Gini and Polarization indexes.

Note: In the left panel is a map, where regions with positive changes in the polarization index are marked with large dots. The upper part of the left panel is the density of changes in polarization index. The lower panel shows the density of changes in the Gini index superimposed on the conditional density of density of changes in Gini index (given that the changes in polarization are positive).